Discrete Mathematics 47 (1983) 117–120
North-Holland

NOTE

A QUINTESSENTIAL PROOF OF VAN DER WAERDEN'S THEOREM ON ARITHMETIC PROGRESSIONS

George MILLS
Carleton College, Northfield, MN 55057, USA

Received 29 October 1982
Revised 26 January 1983 and 15 March 1983

A particularly well suited induction hypothesis is employed to give a short and relatively direct formulation of van der Waerden's argument which establishes that for any partition of the natural numbers into two classes, one of the classes contains arbitrarily long arithmetic progressions.

All currently known proofs of van der Waerden's classical theorem on arithmetic progressions rely on a strengthened induction hypothesis which ends up proving considerably more than the simplest version of the theorem asserts. That this is so stems from the fact that these proofs all are based on the same key idea introduced by van der Waerden back in 1927 [4]. The proof offered in this paper is also based on this idea, but it is formulated as simply and directly as possible, so as to exhibit most clearly the quintessence of van der Waerden's argument without unnecessary details or auxiliary concepts. This is accomplished by choosing an induction hypothesis which seems particularly well suited to the nature of the argument.

Several advantages result from this. For one thing it is not necessary to carry out a double induction on two variables. A single induction suffices. Also the number of classes can be left fixed at two throughout the argument. The usual generalization to an arbitrary number of classes is simply a coding trick which can be more elegantly handled by generalizing a different parameter. (However the present proof trivially generalizes to r classes if desired. For a different proof that avoids this generalization, see [1].) The notation in the present proof is quite transparent (after the definition of the key term 'd, s-translatable' has been mastered). For example there is no need for double subscripts. Finally the proof provides simple, explicit recursion equations for the upper bound function associated with the theorem. Unfortunately this upper bound still grows as rapidly as Ackermann's function, but that is an inherent feature of van der Waerden's argument which no reformulation can avoid.

A word on notation. All number variables range over nonnegative integers. An (arithmetic) progression of length s is a set of the form $\{a+id \mid 0 \leq i < s\}$ with $d>0$. If $d=1$ this is called a block, and will be denoted $\langle a, s\rangle$. Subprogression means progression which is a subset, and similarly for subblock. We will prove van der Waerden's Theorem in the following relativized, finitary form. (The usual form corresponds to $b=1$ and '$u=\infty$'.)

Theorem (van der Waerden). *For any s there exists u such that for any block $\langle b, u\rangle$ and partition $C:\langle b, u\rangle \to \{0, 1\}$ there exists a subprogression of $\langle b, u\rangle$ of length s on which C is constant.*

As auxiliary concepts for the proof we define the following in the context of such a partition. We say that a block $\langle a, t\rangle$ is *congruent* to, or is a *translate* of, $\langle a', t\rangle$ (written $\langle a, t\rangle \equiv \langle a', t\rangle$) iff they have the same color pattern, that is $C(a+i)=C(a'+i)$ for $0\leq i<t$. This is clearly an equivalence relation. A key observation is that for a given t there are at most 2^t equivalence classes.

To prove van der Waerden's Theorem we will build up not simply progressions of like-colored points but progressions of congruent blocks. Here is the key definition: $\langle a, t\rangle$ is *d, s-translatable* in $\langle b, u\rangle$ iff $d>0$ and

(1) $\langle a+jd, t\rangle \subseteq \langle b, u\rangle$ for $0\leq j\leq s$, and

(2) $\langle a, t\rangle \equiv \langle a+jd, t\rangle$ for $0\leq j<s$ (not necessarily for $j=s$).

That is, in $\langle b, u\rangle$ there are s congruent translates of $\langle a, t\rangle$, spaced d apart, with room for one more. Of course the only constraints put by (1) are that $a\in\langle b, u\rangle$ and $a+sd+t\leq b+u$. Observe that any subblock of $\langle a, t\rangle$ will then also be d, s-translatable in $\langle b, u\rangle$.

Now define

$W_s(t)=W(s, t)=$ smallest u such that for any $C:\langle b, u\rangle \to \{0, 1\}$ there exist a, d such that $\langle a, t\rangle$ is d, s-translatable in $\langle b, u\rangle$.

Van der Waerden's Theorem is equivalent to asserting that $W(s, 1)$ exists for all s, since $\langle a, 1\rangle=\{a\}$, giving us the progression $a, a+d, a+2d, \ldots, a+(s-1)d$ of like-colored points. The proof of the thoerem, however, will show that for all s $W(s, t)$ exists for all t, and this will be done by induction on s.

First observe that $W(1, t)=t+1$, since condition (2) is vacuous in this case: let $a=b$ and $d=1$. Now assume $s\geq 1$ and $W_s(t)$ exists for all t. Then we have the following explicit upper bound for $W_{s+1}(t)$ for any given t, proving that $W_{s+1}(t)$ exists.

Lemma. $W_{s+1}(t)\leq 2W_s^{(n)}(t)$, *the nth iterate of W_s, where $n=2^t$.*

Proof. Let $u=2W_s^{(n)}(t)$ and $C:\langle b, u\rangle \to \{0, 1\}$. Define $t_0=t$ and $t_{i+1}=W_s(t_i)$, so $t_n=\frac{1}{2}u$. Let $P_n=\langle b, t_n\rangle$ and by 'reverse recursion' choose blocks $P_i=\langle a_i, t_i\rangle$ and integers $d_i>0$ for $i=n-1, n-2, \ldots, 2, 1, 0$ so that P_i is d_i, s-translatable in P_{i+1}.

Now for $0 \leq i \leq n$ let $b_i = a_0 + s(d_0 + d_1 + \cdots + d_{i-1})$. By the pigeonhole principle and our remark on the number of equivalence classes, there must exist $0 \leq p < q \leq n = 2^t$ such that $\langle b_p, t\rangle \equiv \langle b_q, t\rangle$. Then $\langle b_p, t\rangle$ is the desired d, $(s+1)$-translatable block in $\langle b, u\rangle$, where $d = d_p + d_{p+1} + \cdots + d_{q-1}$.

To verify this consider first any j with $0 \leq j < s$. Using the fact that P_i (and hence any subblock of P_i) is d_i, s-translatable in P_{i+1} get inductively that

$$\begin{aligned}\langle b_0, t\rangle = P_0 &\Rightarrow \langle b_0 + sd_0, t\rangle \subseteq P_1 \Rightarrow \cdots \langle b_p, t\rangle \subseteq P_p\\ &\Rightarrow \langle b_p + jd_p, t\rangle \subseteq P_{p+1} \Rightarrow \cdots\\ &\Rightarrow \langle b_p + jd_p + \cdots + jd_{q-1}, t\rangle = \langle b_p + jd, t\rangle \subseteq P_q \subseteq \langle b, \tfrac{1}{2}u\rangle\end{aligned}$$

and $\langle b_p, t\rangle \equiv \langle b_p + jd_p, t\rangle \equiv \cdots \equiv \langle b_p + jd, t\rangle$. For $j = s$ the inclusions still hold, but it is the coincidence that $\langle b_p + sd, t\rangle = \langle b_q, t\rangle$ which guarantees $\langle b_p, t\rangle \equiv \langle b_p + sd, t\rangle$. Finally for $j = s+1$ since $b < b_p + sd + t \leq b + \frac{1}{2}u$ it follows that $d < \frac{1}{2}u$, whence $b_p + (s+1)d + t \leq b + u$, whence $\langle b_p + (s+1)d, t\rangle \subseteq \langle b, u\rangle$ as required. □

Note that this proof can be adapted to partitions into $r > 2$ classes simply by changing $n = 2^t$ to $n = r^t$.

With just a minor modification in the basis step, the same construction actually proves a stronger theorem, namely the Extended Hales–Jewett Theorem (see [3, pp. 34–38]). The n-cube over m elements is the set

$$C_m^n = \{(x_0, x_1, \ldots, x_{n-1}) \colon x_i \in \{0, 1, \ldots, m-1\}\}$$

which may be identified with $\langle 0, m^n\rangle$ by the correspondence

$$(x_0, \ldots, x_{n-1}) \leftrightarrow x = \sum_{i=0}^{n-1} x_i m^i.$$

Define $E(x) = \{i : x_i \neq 0\}$. Then a k-dimensional subspace of C_m^n is neither more nor less than a subset of $\langle 0, m^n\rangle$ of the form

$$\{a + y_0 d_0 + y_1 d_1 + \cdots + y_{k-1} d_{k-1} : y_i \in \{0, \ldots, m-1\}\}$$

such that each d_i is a sum of distinct powers of m (i.e., $d_i = \sum m^j$, summing over $j \in E(d_i)$) and the sets $E(a), E(d_0), \ldots, E(d_{k-1})$ are disjoint.

Theorem (Extended Hales–Jewett). *For all m, n there exists N such that if the points of C_m^N are 2-colored there exists a monochromatic n-dimensional subspace.*

This will be a consequence of the following family of assertions ($m \geq 2$ is fixed).

HJ(s): For all t there exists $u = U_s(t)$ such that for any $C : \langle b, u\rangle \to \{0, 1\}$ there exist a, d such that

(1) $\langle a, t\rangle$ is d, s-translatable in $\langle b, u\rangle$,
(2) $d = \sum m^j$, summing over $j \in E(d)$,
(3) $t \leq m^j$ for all $j \in E(d)$, and
(4) $E(a') \cap E(d) = \emptyset$ for all $a' \in \langle a, t\rangle$.

The proof is by induction on s. For $s=1$ we may take $u=(m^2+m+1)t$ by the following argument. Given $C:\langle b, u\rangle \to \{0, 1\}$ choose k and c so that $d=m^c$ and $a=m^{c+1}k$ satisfy $t \leq d < mt$ and $b \leq a < b+m^{c+1}$. Then $E(d)=\{c\}$, so (2) and (3) hold. Also $i \in E(a) \Rightarrow i \geq c+1$, and $j<t \Rightarrow [E(j) \cap E(d)=\emptyset$ and $E(a+j)=E(a) \cup E(j)]$, so (4) holds. Finally

$$a+d+t<(b+m^{c+1})+mt+t<b+m^2t+mt+t=b+u$$

so the nearly vacuous (1) holds.

For the step from s to $s+1$ use exactly the same construction as in the proof of van der Waerden's Theorem to establish $U_{s+1}(t) \leq 2U_s^{(n)}(t)$ where $n=2^t$. The inductive verification of (2), (3) and (4) is easy once one notes that if $0 \leq j<k<n$, $i \in E(d_j)$ and $i' \in E(d_k)$, then $t_j \leq m^i \leq d_j < t_k \leq m^{i'}$. Thus $E(d_p)$, $E(d_{p+1}), \ldots, E(d_{q-1})$ are disjoint, with $E(d)$ being their union, so (2) and (3) hold. Also $a' \in \langle b_p, t\rangle$ implies $a' \in P_j$ for $j=p, p+1, \ldots, q-1$, so $E(a') \cap E(d_j)=\emptyset$ for such j, and (4) follows.

Now to prove the Extended Hales–Jewett Theorem, given m, n choose N so that $m^N \geq U_m^{(n)}(1)$. Then given a 2-coloring of C_m^N, identify it with a map $C:\langle 0, U_m^{(n)}(1)\rangle \to \{0, 1\}$ and extract a nested tower of blocks $\{a\}=\langle a_0, t_0\rangle \subseteq \langle a_1, t_1\rangle \subseteq \cdots \subseteq \langle a_n, t_n\rangle$ so that $\langle a_i, t_i\rangle$ is d_i, m-translatable in $\langle a_{i+1}, t_{i+1}\rangle$ and the conditions corresponding to (2), (3) and (4) hold. Then a, $d_0, d_1, \ldots, d_{n-1}$ give us our monochromatic n-dimensional subspace of C_m^N for the same reason that the induction hypothesis is preserved in the construction from HJ(s) to HJ($s+1$). □

Acknowledgements

This formulation of van der Waerden's argument was inspired by the paper [2] and by Ron Graham's elegant exposition (reproduced in [3]) of van der Waerden's Theorem at the C.B.M.S. Regional Conference on Ramsey Theory, St. Olaf College, June, 1979. I also wish to thank Tom Savage for encouraging me to simplify the known proofs. The centrality and sufficiency of the concept d, s-translatable became apparent to me when I formulated the argument in terms of a nonstandard model of arithmetic with u an infinite integer.

References

[1] W. Deuber, On van der Waerden's theorem on arithmetic progressions, J. Combin. Theory (A) 32 (1982) 115–118.
[2] R.L. Graham and B.L. Rothschild, A short proof of Van der Waerden's theorem on arithmetic progressions, Proc. Amer. Math. Soc. 42 (1974) 385–386.
[3] R.L. Graham, B.L. Rothschild and J. Spencer, Ramsey Theory, (Wiley, New York, 1980).
[4] B.L. van der Waerden, Beweis einer Baudetschen Vermutung, Nieuw Arch. Wisk. 15 (1927) 212–216.